

OM NPs ISFREKVENSPRODUKT – INNGANGSDATA OG METODIKK

Dag Vongraven og Mikhail Itkin
Norsk Polarinstitut
25. september 2018

Hva er isfrekvens?

Isfrekvens er et enkelt statistisk mål for forekomst av is, bygd på historiske data. Isfrekvens tilsvarer det som i engelsk fagterminologi heter «ice persistence».

Hvorfor isfrekvens?

Isfrekvens bruker historiske data til å si noe om sannsynligheten for is på kort sikt. Hvor godt dette reflekterer hva som kan forventes avhenger av dataperiodens lengde, hvorvidt endringene i isforhold skjer raskt og om endringene forventes fortsatt å være raske.

Isfrekvens benyttes også fordi det er den best tilgjengelige angivelsen av hvor iskantssonen sannsynligvis vil befinne seg i en gitt måned. Iskantsonen er svært dynamisk, og endres og forflyttes kontinuerlig, men den befinner seg med 100% sikkerhet mellom grensene for 0 og 100% isfrekvens, dvs mellom grensene for hvor det i en gitt måned alltid er is og aldri er is. Isfrekvens sier altså noe om hvor sannsynlig det er at is skal forekomme en gitt måned, noe andre isparametere ikke gjør, som f.eks. isdekke eller iskonsentrasjon.

Hvilke historiske data benyttes og hvor kommer de fra?

Rådata i form av daglige målinger av iskonsentrasjonen kommer fra National Snow and Ice Data Center (NSIDC) i Boulder, Colorado, USA. Dette datasettet viser iskonsentrasjon over hele Arktis, målt fra satellitter. Iskonsentrasjon måles ved at satellittinstrumentet registrerer varmestråling fra havoverflaten (mikrobølger). Den registrerte strålingsmengden avhenger først og fremst av hvor mye vannoverflate som er dekt av is, fordi isen fungerer som varmeisolasjon. Med utgangspunkt i dette kan vi beregne iskonsentrasjonen. Om våren og sommeren kan beregningene bli påvirket av at isoverflaten er delvis dekt av smeltevann eller våt snø. Dette kan gi en usikkerhet på opptil 20-30 %. Om vinteren er feilmarginen betydelig mindre, i gjennomsnitt 7 %. Vann og ispartikler i atmosfæren kan også påvirke satellittens målinger.

Rådata lastes ned som «rasterdata». Til forskjell fra vektordata, hvor data endres kontinuerlig mellom to punkter på kartet, gir rasterdata enkeltverdier for hver rute i et rutenett. NSIDC gir oss iskonsentrasjonsdata i et 25 x 25 km rutenett. En 25 x 25 km rute kalles et piksel. Hvert piksel har en verdi for iskonsentrasjon for hver dag. Når vi er ute etter å beregne isfrekvens sier vi at det innenfor et piksel en gitt dag er is hvis iskonsentrasjonen for denne dagen er 15% eller større, og ikke is hvis iskonsentrasjonen er mindre enn 15%. 15% er her den såkalte terskelverdien, som også kan kalles en deteksjonsgrense (dvs. for registrering av is ved hjelp av fjernmåling). Bruk av 15% som terskelverdi er en global standard innen relevante fagfelt.

De opprinnelige rasterdataene for iskonsentrasjon blir deretter behandlet på Norsk Polarinstitutt (NP) hvor man beregner isfrekvens for hvert piksel. Man velger en dataperiode, som for disse isfrekvensprodukter er satt til de foregående 30 årene. Tretti år er valgt som dataperiode fordi dette er standard normalperiode innen klimatologi. NP beregner isfrekvens måned for måned innenfor denne 30-års dataperioden.

Det finnes datasett med bedre oppløsning, men det er flere grunner til at vi benytter data fra NSIDC. Hovedårsakene er at dataserien er konsistent over hele perioden fra 1978 til i dag, dvs målt med samme metodikk over hele perioden, samt at dataserien er svært godt dokumentert.

Isfrekvens og isdekke

Det er helt essensielt å være klar over forskjellen mellom isfrekvens og isdekke.

Prosent isfrekvens sier noe om andelen av dager i en dataperioden hvor det forekommer is med iskonsentrasjon > 15%. Her er det viktig å huske at isfrekvens ikke sier noen ting om hvilken spesifikk iskonsentrasjon det er innenfor hvert piksel, bare at det er mindre eller mer enn 15% på en gitt dato.

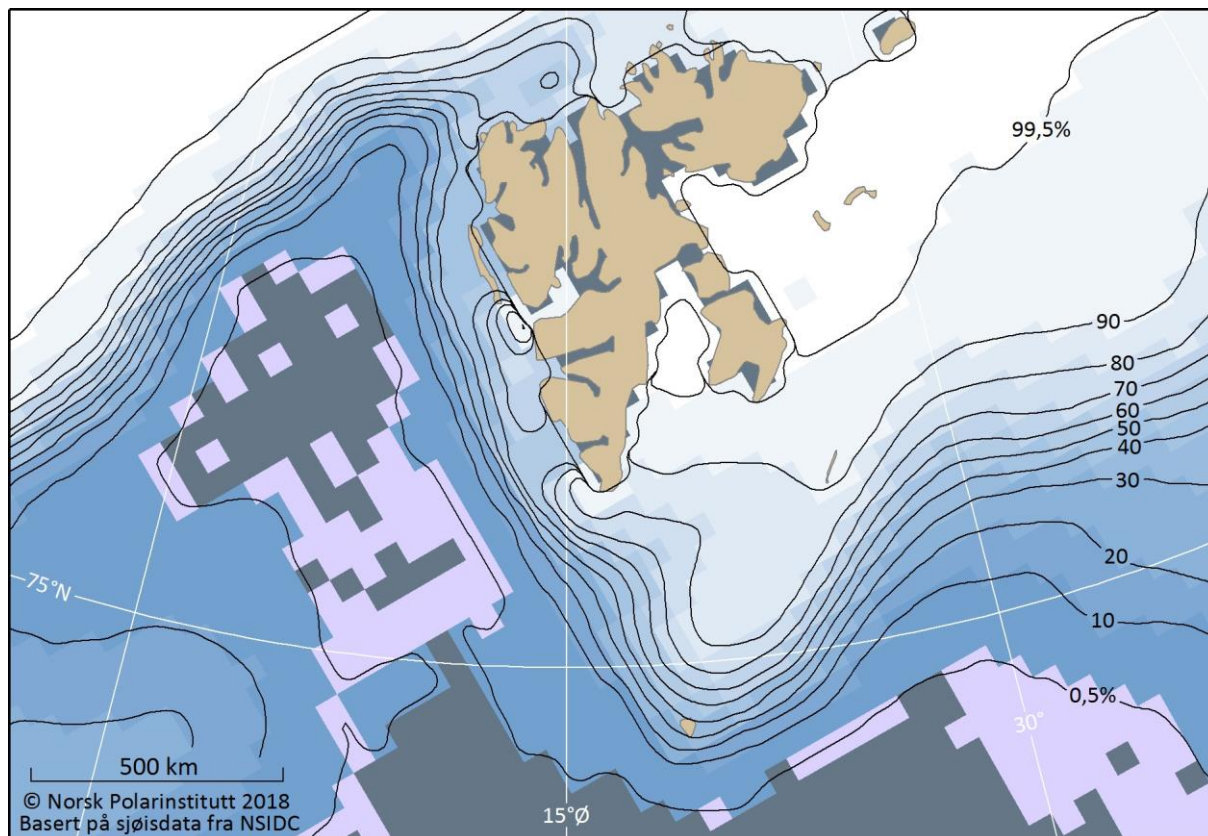
Prosent isdekke sier noe om andelen av et areal som er dekket av is. Isdekket angis i prosent eller kvadratkilometer.

0 og 100% isfrekvens

Som et eksempel, hvis vi har en isfrekvens på 48% i et gitt piksel, så vil det bety at det for den aktuelle måneden har vært is (>15% iskonsentrasjon) i dette pikselet i 48% av det totale antall dager i dataperioden, som i NPs produkt er 30 år. En isfrekvens på 0% betyr at det aldri har vært is i dette pikselet i den aktuelle måneden, og en isfrekvens på 100% betyr at det har vært is der hver eneste dag i den aktuelle måneden i en gitt dataperiode.

Grensen for 0% isfrekvens markerer også det som kalles maksimum isutbredelse i gjeldende dataperiode, og 100% isfrekvens det som kalles minimum isutbredelse. Sagt på en annen måte så er maksimal isutbredelse yttergrensen for det største aggregerte arealet med isfrekvensverdier større enn 0%.

Og, selv om rådata eksisterer som et raster, altså diskontinuerlige data i et 25 x 25 km rutenett, så genererer NPs beregninger jevne linjer for 0% og 100% isfrekvens, og for hver 10^{de} % mellom 0 og 100 (se Figur 1).



Figur 1 Figuren viser rasterdata og beregnede glatte linjer. De gule gridcellene er gridceller med verdier > 0%, og de av cellene som ligger utenfor 0.5%-linjen blir filtrert bort under bearbeidelsen av data. Data er fra mars måned.

Hva er nytt i det nye datasettet?

Da NPs isfrekvensprodukt (kart og datasett) ble lansert og først publisert våren 2014 var det med data for 30-årsperioden 1984-2013. Ved lansering var målsetningen å oppdatere det årlig i forbindelse med at NSIDC frigir foregående års datasett for iskonsentrasjon. Datasettet har, i henhold til opprinnelige planer, blitt oppdatert årlig. Siste oppdatering skjer nå høsten 2018 med data for perioden 1987-2016. Data for 2017 har enda ikke blitt tilgjengeliggjort.

I forbindelse med publiseringen av 2016-data har NSIDC gjort noen endringer i sine metoder. Fjernmålingene registrerer som tidligere nevnt mikrobølgestråling fra havoverflaten, som dempes av is eller vannpartikler. NSIDC benytter nå to algoritmer¹ for å få fram det beste datasettet for iskonsentrasjon, i motsetning til tidligere, hvor de kun benyttet en av dem. Det er også tatt i bruk en ny prosedyre for automatisert kvalitetskontroll av datasettet. Denne prosedyren sørger bl.a. for å filtrere ut støy som stammer fra væreffekter, f.eks. fra store vannpartikler i atmosfæren som misvisende kan registreres som is, og støy som skyldes overgangen mellom land og hav. NSIDC benytter også forbedrete maskering for å fjerne ismålinger i områder hvor is aldri forekommer. Kombinasjonen av to algoritmer gjør at datasettet har bedre kvalitet enn tidligere, og det betyr også at all nødvendig etterbearbeiding av data nå gjøres av NSIDC.

¹ En algoritme er et sett av regneregler som benyttes ved matematiske beregninger eller annen problemløsning, oftest benyttet i relasjon til datamaskiner.

0.5 % isfrekvens

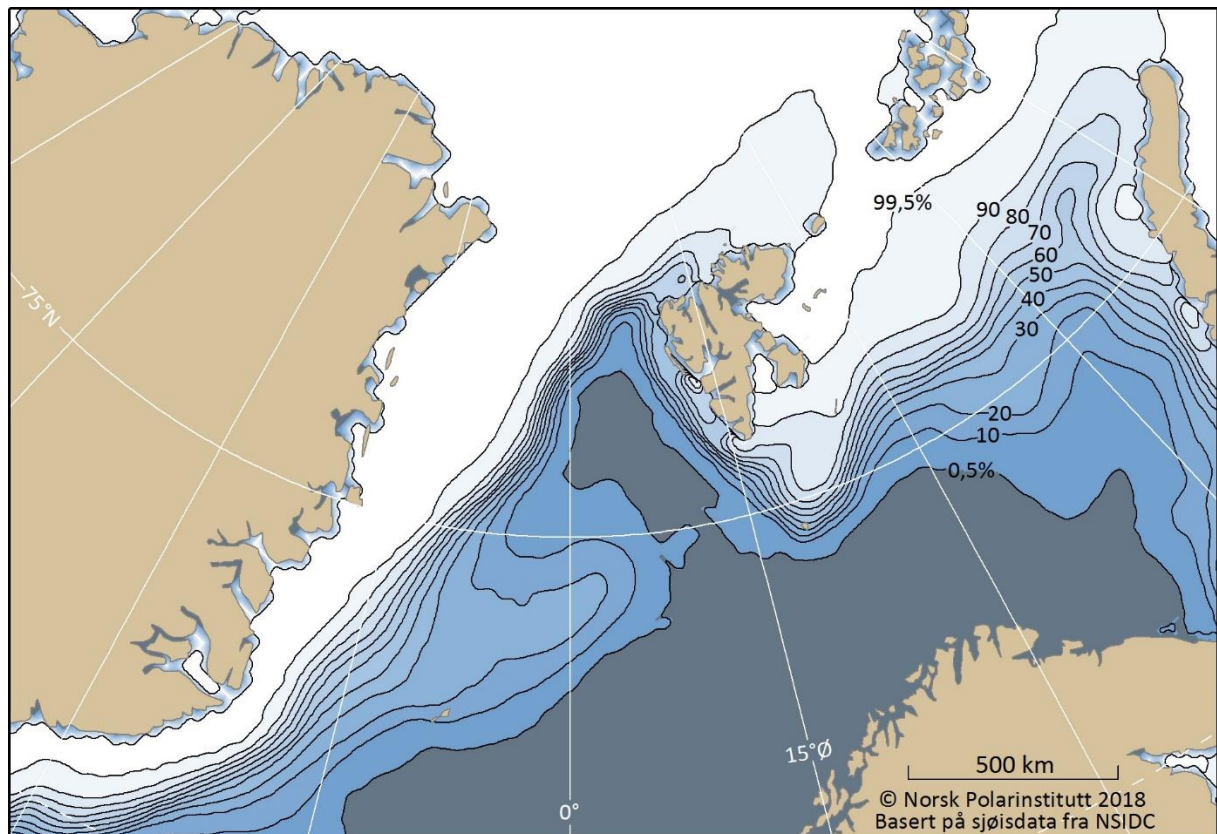
Selv om kvaliteten på rådata fra NSIDC er forbedret, vil rådata alltid inneholde støy, som alle andre typer rådata. Støykildene er mange, bl.a. måleinstrumenter og beregningsalgoritmer. For eksempel vil beregning fra rasterdata til kontinuerlige jevne linjer kreve bruk av algoritmer som kan generere støy. Støy i denne sammenhengen vil f.eks. kunne være at de endelige dataene viser is i områder der det ikke har vært is. For å begrense hvor mye slik støy som vises på våre kart, så har NP valgt å ikke inkludere isfrekvens lavere enn 0.5%. Denne grenseverdien er satt så lavt som mulig, slik at vi ikke risikerer å fjerne reelle isforekomster sammen med støyen. Dette gjelder imidlertid kun de beregnede glattede linjene som vises i kartet (se Figur 2). Rasterdata vil fortsatt vise verdier lavere enn 0.5%.

Datasettet

Datasettet kan lastes ned fra NPs dataservert:

<https://data.npolar.no/dataset/a89b2682-e1f8-44b5-ab73-e6ba65f3a7a7>

Datasettet består av både rasterdata (rutenett) og shapefiler (kontinuerlige linjer). Shapefilene er laget for hver 10^{de} % isfrekvens.



Figur 2 Oversiktskart over NPs isfrekvensprodukt, med verdier lavere enn 0.5% isfrekvens utelatt, og data fra perioden 1987-2016. Data er fra mars måned.